DOCUMENT RESUME

ED 277 740 TM 870 017

AUTHOR Merrill, Beverly; Peterson, Sarah
TITLE Writing Sample Assessment: Reliable, Efficient, and Computerized.
PUB DATE 20 Apr 86
NOTE 7p.; Paper presented at the Annual Meeting of the American Educational Research Association (67th, San Francisco, CA, April 16-20, 1986).
PUB TYPE Speeches/Conference Papers (150) -- Reports - Descriptive (141)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Computer Assisted Testing; Elementary Secondary Education; Graduation Requirements; *Holistic Evaluation; *Interrater Reliability; Minimum Competency Testing; *Scoring; Testing Problems; *Writing Evaluation
IDENTIFIERS *Mesa Public Schools AZ

ABSTRACT

When the Mesa, Arizona Public Schools initiated an ambitious writing instruction program in 1978, two assessments based on student writing samples were developed. The first is based on a ninth grade proficiency test. If the student does not pass the test, high school remediation is provided. After 1987, students must pass this test in order to graduate from high school. The more ambitious assessment is conducted twice per year in grades 1 through 6. Students write a short composition on a prescribed topic. The writing samples are collected, coded, and holistically scored. This evaluation is used primarily to determine student growth and to assess the writing programs. Procedures established to ensure anonymity and reliability in grading the writing samples became tedious to administer. However, in spite of the clerical workload, the writing assessment program persisted. Last year, a computerized scoring process was developed. Each writing sample is read and scored twice, with an overall interrater reliability of .68. Coded scores are recorded on the computer sheet, and fall and spring data are consolidated by the data processing department. The computer system identifies raters who need additional inservice education to improve reliability. Teachers are pleased with the computerized system. (GDC)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED277740

Writing Sample Assessment: Reliable, Efficient and Computerized

Beverly Merrill
Sarah Peterson
Mesa Public Schools
Mesa, Arizona

In 1978, under the direction of the governing board, Mesa Public Schools initiated a very ambitious writing program. Since that time, the improvement of student writing skills has received special attention throughout the district. Writing is viewed as a means of teaching thinking, as a tool for learning, and as an essential job skill. As the writing program evolved, the following components emerged: a language arts curriculum program with a writing instruction focus, a district-wide writing-across-the-curriculum program, an extensive school and district level inservice program which includes a university credit class designed and offered solely for district educators, classroom demonstration teaching, an intra-district pen pal program, a highly competitive writing contest, and two assessment programs based on student writing samples. All of this is done in support of improved classroom instruction with the writing process philosophy underscoring all components of the program.

Paper presented at AERA convention, San Francisco, April 20, 1986.

TM 870 017

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

B. Merrill

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

One of the assessment programs is based on proficiency. At the conclusion of each year, ninth grade students are required to take a test of their writing skills. This writing sample is evaluated by a team of carefully trained teacher-raters to determine if ninth grade students have reached writing proficiency at that grade level. If a student does not pass the proficiency exam, then remediation takes place in the high schools. Starting with the graduating class of 1987, students must demonstrate writing proficiency at the ninth grade level before they can graduate from a Mesa high school.

The second assessment, and the more ambitious of the two, is conducted twice a year in grades one through six. At each of the district's elementary schools, in the fall and in the spring, students are asked to write a short composition on a prescribed topic. These writing samples are collected, coded and holistically scored (the paper's total impact--general impression marking) at an inservice day in the spring. This evaluation system is used primarily to determine growth. Information from the writing sample data helps assess the writing programs at the classroom, school and district levels. All teachers participate in the evaluation, a reflection of the writing program philosophy which states that all educators, regardless of subject-area, are responsible for fostering language development.

The assessment at the elementary schools has a long and interesting history. At first, teachers had to be assured that the results of the writing sample evaluation would be accurate.

Ensuring reliable evaluations means the process has to adhere to the following tenents: the papers are always read anonymously, the two raters who score the papers do so independently (third readers are called in if the first two scores are not adjacent), and the raters are trained in how to establish a mind-set in order to make the scoring consistent. It was the independent scoring by more than one rater that made the process difficult. Paper codes were created, copied, and matched; scores were assigned, added, and reassembled. This was all done by hand. Soon coded bundle tallies began to accumulate and reproduce in shoe boxes. In the very early stages of the testing program, teachers in the Mesa system had to endure some very awkward assessments.

It is not difficult to understand why some school districts, although philosophically committed to testing via writing samples, have a difficult time justifying the process. In pre-computerized days, the time spent on the actual reading of the students' papers paled in contrast to the time spent on the managerial part of the process. Chief readers, who oversee the procedure, report that at one time approximately thirty to forty hours were spent preparing the papers for scoring and approximately forty-five to fifty hours were spent assembling and recording the scores after the evaluations were completed. This part of the process became increasingly embarrassing and was not easily defended.

Despite the clerical unwieldyness, however, MPS clung tenaciously to writing sample assessment. If writing is

tested, it will be taught. And, there were very important tangential benefits. During the mind-set procedure, teachers have a unique opportunity to discuss writing. Their experiences in the scoring sessions help them gain insight into the nature of the writing process. The rubric calls for the various components of writing to be considered simultaneously and this helps teachers view writing as an integrative process. Participants learn that they can agree on the scores they choose to assign to the papers and that they can justify those scores according to the standards they developed in the mind-set group.

Last year a computerized scoring process was developed which has all but eliminated the clerical procedures and made it possible to analyze the writing sample data in a variety of ways. For each sample, computerized sheets are pre-slugged with student number, school, teacher number, and grade level. All of this is on the same sheet that the student uses for the final draft composition. The anonynmity of each writing sample is ensured and (hurray!) there is no transferring of scores. Raters are provided with a rubric, topic and grade level information, but all other identification--including whether it is a fall or a spring sample--is coded. Each writing sample is read and scored twice and the coded scores are recorded on the computer sheet. The scored writing samples are then returned to the data processing department where the information is consolidated. The fall and spring scores are compared to determine if there has been growth in writing abilities during

the school year. The data from the evaluation provides a systematic way of assessing the writing program.

The 1984-85 assessment showed the inter-rater abilities were as follows:

Fall		Spring	
Grade	Reliability	Grade	Reliability
1	.75	1	.73
2	.63	2	.67
3	.63	3	.66
4	.66	4	.64
5	.60	5	.60
6	.61	6	.63
Average:	**.70**	**Average:**	**.66**

Reliabilities for each of the elementary schools in the district ranged from .60 to .75. These are moderately high correlations with an overall reliability score of .68. It is interesting to note that there is less correlation in the upper grades than in the primary grades. Perhaps this is because the primary papers are less complex and are, therefore, easier to assess. With the computerized scoring sheets, it can be determined which raters need additional inservice. Research on inter-rater reliability conducted by P.B. Diederich of Educational Testing Services suggests that the crucial factors in high and low inter-rater reliability are the quality of the training and the educational homogeneity of the raters. According to Dieterich, subjectivity can be reduced by

carefully training raters who have similar educational backgrounds.[1]

Although a formal evaluation was not completed to determine teachers' attitudes toward the computerized process, informal testimony confirms the general notion that teachers are exceedingly pleased to have a process that saves so much time. And, somehow, the use of the main frame makes the whole procedure appear ever so much more scientific, less subjective, more accurate, and, therefore, less threatening. The computerized process is a boon to all who are involved in the MPS writing sample assessment.

Now in its eighth year, the Mesa Public Schools writing program with all its facets is one that the district is fully committed to continuing and expanding. Writing assessment will continue to be a valuable part of this program.

---

1. Paul B. Diederich, Measuring Growth in English. National Council of Teachers of English, Urbana, 1974, p.33.